

RESEARCH ARTICLE

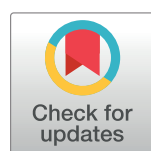
# Identification of C-T novel polymorphism in 3<sup>rd</sup> exon of *OsSPL14* gene governing seed sequence in rice

Muhammad Salah ud Din<sup>1</sup>\*, Xiukang Wang<sup>2\*</sup>, Salman Alamery<sup>3</sup>, Sajid Fiaz<sup>4\*</sup>, Haroon Rasheed<sup>5,6</sup>, Muhammad Abid Khan<sup>5\*</sup>, Shahid Ullah Khan<sup>7</sup>, Sumbul Saeed<sup>8</sup>, Niaz Ali<sup>1</sup>, Kalim Ullah Marwat<sup>9</sup>, Kotb Attia<sup>3</sup>, Itoh Kimiko<sup>10</sup>, Shabir Hussain Wani<sup>11</sup>

**1** Department of Botany, Hazara University, Mansehra, Khyber Pakhtunkhwa, Pakistan, **2** College of Life Sciences, Yan'an University, Yan'an, Shaanxi, China, **3** Department of Biochemistry, College of Science, King Saud University, Riyadh, Saudi Arabia, **4** Department of Plant Breeding and Genetics, The University of Haripur, Haripur, Pakistan, **5** Department of Botany, University of Science and Technology Bannu, Khyber Pakhtunkhwa, Pakistan, **6** Institute of Nuclear Agriculture Sciences, College of Agriculture and Biotechnology, Zhejiang University, Zijingang Campus, Hangzhou, P.R. China, **7** Department of Biochemistry, Women Medical and Dental College Abbotabad, Khyber Pakhtunkhwa, Pakistan, **8** College of Plant Sciences and Technology, Huazhong Agricultural University, Wuhan, P.R. China, **9** Department of Zoology, Kohat university of Science and Technology, Kohat, Khyber Pakhtunkhwa, **10** Institute of Science and Technology, Niigata University, Nishi-ku, Niigata, Japan, **11** Mountain Research Center for Field Crops, Khudwani, Sher-e-Kashmir University of Agricultural Sciences and Technology Srinagar, Srinagar, J&K, India

\* These authors contributed equally to this work.

\* wangxiukang@yau.edu.cn (XW); sfiaz@uoh.edu.pk, akhan5408@gmail.com (SF); abidkhan03@yahoo.com (MAK)



## OPEN ACCESS

**Citation:** Din MSu, Wang X, Alamery S, Fiaz S, Rasheed H, Khan MA, et al. (2022) Identification of C-T novel polymorphism in 3<sup>rd</sup> exon of *OsSPL14* gene governing seed sequence in rice. PLoS ONE 17(3): e0264478. <https://doi.org/10.1371/journal.pone.0264478>

**Editor:** Umair Ashraf, University of Education Lahore, Pakistan, PAKISTAN

**Received:** October 30, 2021

**Accepted:** February 10, 2022

**Published:** March 14, 2022

**Copyright:** © 2022 Din et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting information files](#).

**Funding:** The publication of the present work is supported by the Natural Science Basic Research Program of Shaanxi Province (grant no. 2018JQ5218) and the National Natural Science Foundation of China (51809224), Top Young Talents of Shaanxi Special Support Program. The funders had no role in study design, data collection

## Abstract

Recently food shortage has become the major flagging scenario around the globe. To resolve this challenge, there is dire need to significantly increase crop productivity per unit area. In the present study, 24 genotypes of rice were grown in pots to assess their tillering number, number of primary and secondary branches per panicle, number of grains per panicle, number of grains per plant, and grain yield, respectively. In addition, the potential function of miR156 was analyzed, regulating seed sequence in rice. Furthermore, *OsSPL14* gene for miR156 was sequenced to identify additional mutations within studied region. The results demonstrated Bas-370 and L-77 showed highest and lowest tillers, respectively. Bas-370, Rachna basmati, Bas-2000, and Kashmir Basmati showed high panicle branches whereas, L-77, L-46, Dilrosh, L-48, and L-20 displayed lowest panicle branches. Bas-370 and four other studied accessions contained C allele whereas, L-77 and 18 other investigated accessions had heterozygous (C and T) alleles in their promoter region. C-T allelic mutation was found in 3<sup>rd</sup> exon of the *OsSPL14* gene. The sequence analysis of 12 accessions revealed a novel mutation (C-T) present ~2bp upstream and substitution of C-A allele. However, no significant correlation for novel mutation was found for tillering and panicle branches in studied rice accessions. Taken together present results suggested novel insight into the binding of miR156 to detected mutation found in 3<sup>rd</sup> exon of the *OsSPL14* gene. Nevertheless, L-77, L-46, Dilrosh, L-48, and L-20 could be used as potential breeding resource for improving panicle architecture contributing yield improvement of rice crop.

and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Rice (*Oryza sativa* L.) is among essential monocotyledonous plants belongs to family *Poaceae* and consumed by more than 50% of people around the world. Global food security is becoming a serious challenge for rapidly growing human population owing to negative environmental conditions caused by climate change [1, 2]. Therefore, it is important to accelerate the productivity of crops through breeding superior varieties. Among various attributes, tillering is considered an essential characteristic to determine rice yield because its significant correlation with cereal production [3]. Plants with high tillers might be useful for adapting to diverse environmental condition however, low tillers severely influence extreme stress conditions [4]. In Pakistan, rice has become next staple food after wheat, and a source of foreign earning through export. From the last several decades, agriculture system of Pakistan has faced severe challenges [5]. In Punjab Gujranwala, Sheikh Khupura, and Sialkot are the main rice producing areas whereas, few districts of Punjab and Sindh are less rice cultivated areas [6]. On global scale, over three billion Asian people meet their caloric requirements from rice [7]. In addition, it has been estimated that from 2001 to 2025, global demand for rice production is expected to increased by 25%, and an annual increase of 5.9 metric tons is required to achieve. However, to meet caloric demand for ever-increasing human population, the Food and Agriculture Association of the United States forecasted that by 2050 the rice demand will exceed by 524 metric tons, and productivity of rice is estimated to increase by 2 metric tons per year compared to current rice production [8].

During the last three years, rice production has reached to a record level. During 1940–1960s, green revolution led to an expansion in creation of agro-based industry especially in developing countries, primarily through progress in research and innovation [9]. The new rice varieties, grown on irrigated land in half of the world's crops, contribute almost three-quarters to total rice production. Fifteen countries are contributing 90% of the world's rice production [10] whereas, only China and India account for 50% of rice cultivation.

Tillering begins around 40 days after planting and can last up to 120 days. It is a physiological process of continuous underground branching of compact node joints of the primary shoot. Tillering gives the crop the necessary number of stalks required for a good production. Tillers emerging in this way from nodes in the main culm are called primary tillers. These emerge all through the vegetative growth phase, but stop when plants reach panicle initiation (the starting point of the reproductive phase). These tillers emerge and branch out from the base of the plant. After panicle initiation, tillers may continue to emerge from preexisting tillers, filling out the space among plants. These tillers are called secondary tillers. Plant hormones such as auxin, cytokinin [2, 11], and strigolactones [12, 13] control axillary bud outgrowths. The main *PROTEIN BINDING TO THE SOUAMOSA 14 PROMOTER* (*OsSPL14*) is regulated by micro RNA Osmi156, the improved level of transcription and protein *OsSPL14* results in an IPA phenotype and greater cereal productivity [2, 14]. The QTL *WFP* (*WEALTHY FARMER'S PANICLE*) encoding *OsSPL14* contrarily manages the numbers of tillers in the asexual phase and exactly switches the quantity of rachis in the reformatory period. The greater manifestation of *OsSPL14* in young panicles stretched panicle wafting and grain number per panicle, contribute increase yield of rice [15, 16].

Lower DNA methylation on this allele increases the transcription frequency of *OsSPL14*, which leads to a high number of grains per panicle. There is a substitution of nucleotides C to T from the start codon in the promoter region at a distance of 4.2 kb. Nucleotide C is a target methylation site; nucleotide T avoids methylation with a higher manifestation of *OsSPL14* [14, 17]. Another allele with synonymous substitution of C for A at 3<sup>rd</sup> exon of *OsSPL14*, a target site (complementary site in the coding region) of OsmiR156. Previous studies reported that

OsmiR156 targets approximately 11 *OsSPL* genes. OsmiR156 cleaves the *OsSPL14* mRNA, if *OsSPL14* carries the wild-type nucleotide C at target site.

OsmiR156 is strongly expressed in vegetative shoots, leaves, and the tip but is suppressed in the early stages of panicle development [14, 18]. Introduction of the high-efficiency *OsSPL14WFP* allele to the standard Nipponbare rice variety leading to increased rice production [14]. In-plant architecture, tillering is an important component that increases rice yield and crop components. As a complex agronomic element, a rice plant's grain yield can be duplicated by three-segment attributes: number of panicles per plant, number of grains per panicle, and grain weight. The quantity of panicles relies upon the plant's capacity to deliver tillers (tillering limit), comprising essential, auxiliary, and tertiary tillers. The size of the grain controls the total grain weight. It is indicated by its three measurements (length, width, and thickness) and the level of substantial [2, 11]. The genes *EHD1* and *HD1* involve in photoperiodic controlling and ultimately tillering. An expanded level of articulation of *HD3A* and *RFT1* reduces the number of essential branches per panicle in the line that consolidates *HD1* and *EHD1* [19]. *HD1* builds up the number of spikelets per panicle and grain by suffocating the *Hd3a* joint and shifting the date. Here we reported the genetic and molecular analysis of the *OsSPL14* gene and proposed some functional features of rice genotypes for the evaluation of *OsSPL14* in panicle branching and the regulation of grain yield in rice. Hence our proposed result may suggest that the new mutation in *OsSPL14* can play a significant and potential role in tillering, and grain yield in rice, respectively.

## Material and methods

The study was conducted at the Department of Botany, University of Science and Technology, Bannu during year 2018. A total of 24 rice genotypes and varieties were investigated in current study (S1 Table). For convenience, all these selected genotypes and varieties are referred to as lines in subsequent text. Seeds of the selected germplasm were obtained from NARC, Islamabad, Pakistan. Two lines Basmati-370 and Line-77 were used as check in present investigation and grown under CRD (completely randomized design).

## Germination of rice seedlings

Seed of the selected 24 rice genotypes (Table 1) were grown in replicates following CRD, mature paddy seeds were harvested from each accession.

## Morphological characterization

The plants were scored for six morphological characters based on International Rice Research Institute's descriptor, Manila Philippines [18]. Total number of tillers was recorded by counting total number of tillers from randomly selected three plants and taken their average value. Number of filled grains in a single panicle from three plants was calculated, and their average value was calculated. Number of grains per plant were calculated selecting three random plants from each line. Thousand-grain weight was recorded using a digital balance. Branches per panicle were recorded selecting three plants from each genotype and averaged. Number of secondary branches per panicle was counted from three randomly selected plants of each line and averaged.

## Molecular characterization

DNA extraction was performed from young leaf tissues via modified CTAB method [20]. To validate the expected DNA fragment gel electrophoresis was performed. During current study,

Table 1. Evaluation of important yield contributing traits within selected germplasm.

Accession	NT	NG	BP	NGP	SBP	GW (kg)
BAS-370	27.66 <sup>A</sup>	187.67 <sup>A</sup>	11.667 <sup>AB</sup>	3753.3 <sup>A</sup>	6.3333 <sup>BCDE</sup>	0.0369 <sup>ABCD</sup>
RACHNA Basmati	27.00 <sup>AB</sup>	123.67 <sup>GH</sup>	11.333 <sup>ABC</sup>	2398.7 <sup>B</sup>	7.3333 <sup>ABC</sup>	0.0653 <sup>A</sup>
Bas-2000	25.000 <sup>BC</sup>	172.00 <sup>B</sup>	12.333 <sup>A</sup>	2486.7 <sup>B</sup>	8.0000 <sup>A</sup>	0.0510 <sup>ABC</sup>
Kashmir Basmati	24.667 <sup>BC</sup>	174.00 <sup>B</sup>	12.000 <sup>AB</sup>	1740.0 <sup>C</sup>	6.6667 <sup>ABCD</sup>	0.0387 <sup>ABCD</sup>
Bas-385	23.333 <sup>CD</sup>	178.33 <sup>B</sup>	12.667 <sup>A</sup>	2445.7 <sup>B</sup>	7.6667 <sup>AB</sup>	0.0503 <sup>ABC</sup>
Fakher-e-Malakand	23.000 <sup>CDE</sup>	160.33 <sup>C</sup>	11.667 <sup>AB</sup>	1603.3 <sup>CD</sup>	5.6667 <sup>DEFG</sup>	0.0357 <sup>ABCD</sup>
Swat II	22.000 <sup>DEF</sup>	142.33 <sup>D</sup>	10.667 <sup>BCD</sup>	1406.3 <sup>DEF</sup>	5.0000 <sup>EFGH</sup>	0.0327 <sup>ABCD</sup>
IR9	21.000 <sup>DEFG</sup>	89.67 <sup>L</sup>	8.667 <sup>EF</sup>	1224.7 <sup>F</sup>	5.0000 <sup>EFGH</sup>	0.0407 <sup>ABCD</sup>
L-146	20.667 <sup>EFGH</sup>	118.67 <sup>H</sup>	8.000 <sup>FG</sup>	1186.7 <sup>F</sup>	4.3333 <sup>GH</sup>	0.0263 <sup>BCD</sup>
NPT 22	20.000 <sup>FGH</sup>	137.67 <sup>DE</sup>	8.333 <sup>FG</sup>	1522.7 <sup>CDE</sup>	5.0000 <sup>EFGH</sup>	0.0370 <sup>ABCD</sup>
KSK-133	19.000 <sup>GH</sup>	159.33 <sup>C</sup>	12.333 <sup>ABC</sup>	660.7 <sup>C</sup>	6.6667 <sup>ABCD</sup>	0.0397 <sup>ABCD</sup>
SWAT-I	19.000 <sup>GH</sup>	128.67 <sup>FG</sup>	8.667 <sup>EF</sup>	1286.7 <sup>E</sup>	5.6667 <sup>DEFG</sup>	0.0340 <sup>ABCD</sup>
Bas p-s 2	18.338 <sup>H</sup>	133.33 <sup>EF</sup>	7.000 <sup>CDE</sup>	1333.3 <sup>EF</sup>	6.0000 <sup>CDEF</sup>	0.0333 <sup>ABCD</sup>
L-31	15.667 <sup>I</sup>	96.33 <sup>JKL</sup>	6.333 <sup>H</sup>	775.3 <sup>G</sup>	3.6667 <sup>H</sup>	0.0183 <sup>CD</sup>
Muskan	15.667 <sup>I</sup>	127.00 <sup>FGH</sup>	8.667 <sup>EF</sup>	1270.0 <sup>G</sup>	4.0000 <sup>H</sup>	0.0310 <sup>ABCD</sup>
Malkhar-346	15.333 <sup>I</sup>	103.00 <sup>JK</sup>	9.333 <sup>DEF</sup>	793.0 <sup>G</sup>	5.6667 <sup>DEFG</sup>	0.0237 <sup>CD</sup>
L-134	14.000 <sup>IJ</sup>	108.00 <sup>I</sup>	8.667 <sup>EF</sup>	756.0 <sup>G</sup>	5.0000 <sup>EFGH</sup>	0.0187 <sup>CD</sup>
KSK-434	13.667 <sup>IJK</sup>	90.00 <sup>L</sup>	10.000 <sup>CDE</sup>	660.7 <sup>G</sup>	6.0000 <sup>CDEF</sup>	0.0163 <sup>CD</sup>
P-5	13.333 <sup>IJK</sup>	106.33 <sup>I</sup>	7.000 <sup>GH</sup>	744.3 <sup>G</sup>	3.6667 <sup>H</sup>	0.0150 <sup>CD</sup>
L-20	12.667 <sup>JKL</sup>	92.67 <sup>L</sup>	8.333 <sup>FG</sup>	588.0 <sup>GH</sup>	4.3333 <sup>GH</sup>	0.0433 <sup>ABCD</sup>
L-48	12.667 <sup>JKL</sup>	94.33 <sup>KL</sup>	8.000 <sup>FG</sup>	660.3 <sup>G</sup>	5.6667 <sup>DEFG</sup>	0.0170 <sup>CD</sup>
Dilrosh	12.333 <sup>JKL</sup>	104.67 <sup>IJ</sup>	8.667 <sup>EF</sup>	625.0 <sup>GH</sup>	5.0000 <sup>EFGH</sup>	0.0180 <sup>CD</sup>
L-46	11.333 <sup>KL</sup>	96.00 <sup>JKL</sup>	8.000 <sup>FG</sup>	544.3 <sup>GH</sup>	4.6667 <sup>FGH</sup>	0.0617 <sup>AB</sup>
Line-77	10.666 <sup>L</sup>	66.33 <sup>M</sup>	8.000 <sup>FG</sup>	376.3 <sup>H</sup>	5.0000 <sup>EFGH</sup>	0.0083 <sup>D</sup>
Grand mean	18.250	124.60	9.5278	1328.6	5.5000	0.0330
Maximum	27.667	187.67	12.667	3753.3	8.0000	0.0653
Minimum	10.667	66.33	6.333	376.3	3.6667	0.0083
CV%	8.67	4.33	8.77	11.93	17.60	66.85

NT = Number of tillers, NG = Number of grains per panicle, BP = Branches per panicle, NGP = Number of grains per plant, SBP = Secondary branches per panicle, GW = Grain weight.

<https://doi.org/10.1371/journal.pone.0264478.t001>

seven primers were used along with four pairs of markers to detect the SNPs C/T and C/A in third exon of *OsPL14* gene (S2 Table). These primers were selected from [13] to amplify and sequencing of *OsPL14* gene promoter.

### PCR amplification and nucleotide sequencing

PCR amplification was carried out using PROMEGA kit Madison Company USA. PCR master mix was prepared for a total of 24 samples having 18 µl for each sample. The individual sample reaction contains 12.24 µl of water 2Mm MgCl<sub>2</sub>, taq buffer of 0.4 µl. 0.2 µl for each forward and reverse primer of 20Mm 0.4 of DNTPS and 0.16 of tag polymerase. Template DNA of 2 µl corresponding to 100ng was used. PCR conditions were set for 40 cycles set at 94°C for 2min, followed by 30 seconds 94°C, 30 sec on 55 °C and 72 °C for 30 sec, and final extension of 72 °C for 10 min. Once the desired PCR products were confirmed, template DNA was sent to tsingke Biotechnology Company, China for sequencing of seed sequence for miR156 third exon of *OsSPL14* gene to detect potential mutation.

## Statistical analysis

The morphological data was analysed using, statistic 8.1 [19]. The sequencing result was first aligned to reference rice gene *OsPL14* (838692 gene ID) available at NCBI (<https://www.ncbi.nlm.nih.gov/>). Crustal W multiple alignment tools were used to carried out multiple sequence alignment to identify potential SNPs. BLAST and I-TASSER were utilized to analyze the effect of SNPs for amino acid substitution on protein structure. The PDB file of the selected model was further subject to ERRAT-2 program to generate overall quality factor for *OsSPL14* protein. Estimated RMSD and TM score and respective was run in ERRAT-2 to generate an overall quality score.

## Results

### Morphological traits results

Morphological evaluation of genotypes showed highest number of tillers for Basmati 370, the lowest was recorded for Line 77. Similarly, the highest grains per panicle were recorded for basmati 370 whereas, Line 77 showed the lowest number of grains per panicle. The highest number of panicles per branches were recorded for Basmati 385, whereas, the lowest were observed in Line 33. Highest number of grains per plant were recorded for Basmati 370 whereas, lowest grains per plant were recorded for Line 77. In addition, Basmati 200 had the highest secondary branches per panicle, while P-5 had the lowest secondary branches per panicle.

The grain weight was also accounted for among the investigated germplasm, the highest grain weight was recorded for Rachna, while the lowest was recorded for Line 77. The maximum value for the variation coefficient was recorded for grain weight, and the number of grains per panicle and revealed the lowest. Importantly, all traits showed significance for scored traits, as represented in (Table 1).

### Correlations among important yield contributing traits

The association of various traits is essential to collect information for quality and yield related traits in rice. To understand the magnitude and nature of the association between the investigated traits i.e., number of tillers per plant, number of grains per panicle, primary branches per panicle, number of grains per plant, secondary branch per panicle, and grain weight were subjected to pearson's correlation, results are shown in (Table 2). Pearson's correlation revealed that the number of tillers per plant showed a significant correlation with the number of grains per plant, followed by several grains per panicle. The tillers per plant also showed a strong and highly significant correlation of 0.7620, 0.6313, and 0.5671 with primary branches per panicle, secondary branches per panicle, and grain weight, respectively.

Furthermore, the number of grains per panicle showed a significant correlation with number of grains per plant followed by primary branches per panicle and secondary branches per panicle. A significant correlation was recorded between the number of grains per panicle and grain weight. Primary branches per panicle also showed a highly significant correlation with secondary branches per panicle, followed by the number of grains per plant. The number of primary branches per panicle also revealed a highly significant correlation with the grain weight. The number of grains per plant also manifested a highly significant correlation with secondary branches per panicle, followed by grain weight. Nonetheless, a significant correlation was recorded between secondary branches per panicle and grain weight.

Table 2. Pearson correlation among the six yield attributing traits.

	NT	NG	PBP	NGP	SBP
NG	0.8120**				
PBP	0.7620**	0.8139**			
NGP	0.8987**	0.8539**	0.7629**		
SBP	0.6313**	0.6269**	0.8713**	0.6691**	
GW	0.5671**	0.4649*	0.5261**	0.5446**	0.5066*

NT = Number of tillers, NG = Number of grains per panicle, PBP = Branches per panicle, NGP = Number of grains per plant, SBP = Secondary branches per panicle, GW = Grain weight

\*significant at  $P < 0.05$

\*\* Significant at  $P < 0.01$

<https://doi.org/10.1371/journal.pone.0264478.t002>

## Molecular genotyping

PCR amplification of *SPL14-04* regains for C-T allele genotyping. Highly quality DNA extracted was used for PCR amplification. The *SPL14-14* region was amplified with CR and TR primers that resulted in 267bp amplicons (Fig 1, Table 3). Out of the 24 rice lines and varieties used, 5 had C type allele, whereas 19 contain heterozygous with a C-T allele. The reference genome of *SPL14* undergo both symmetric and asymmetric methylation therefore, C allele is a target nucleotide for methylation in the promoter region of the *OsSPL14* gene, which regulates the expression level of a gene. The C-T substitution prevents methylation and may interfere with the transcription of gene through epigenetic phenomena. Basmati 370, IR9, Malkhar-346, Line 20, and Line 31 carry wild type C allele while Rachna basmati,

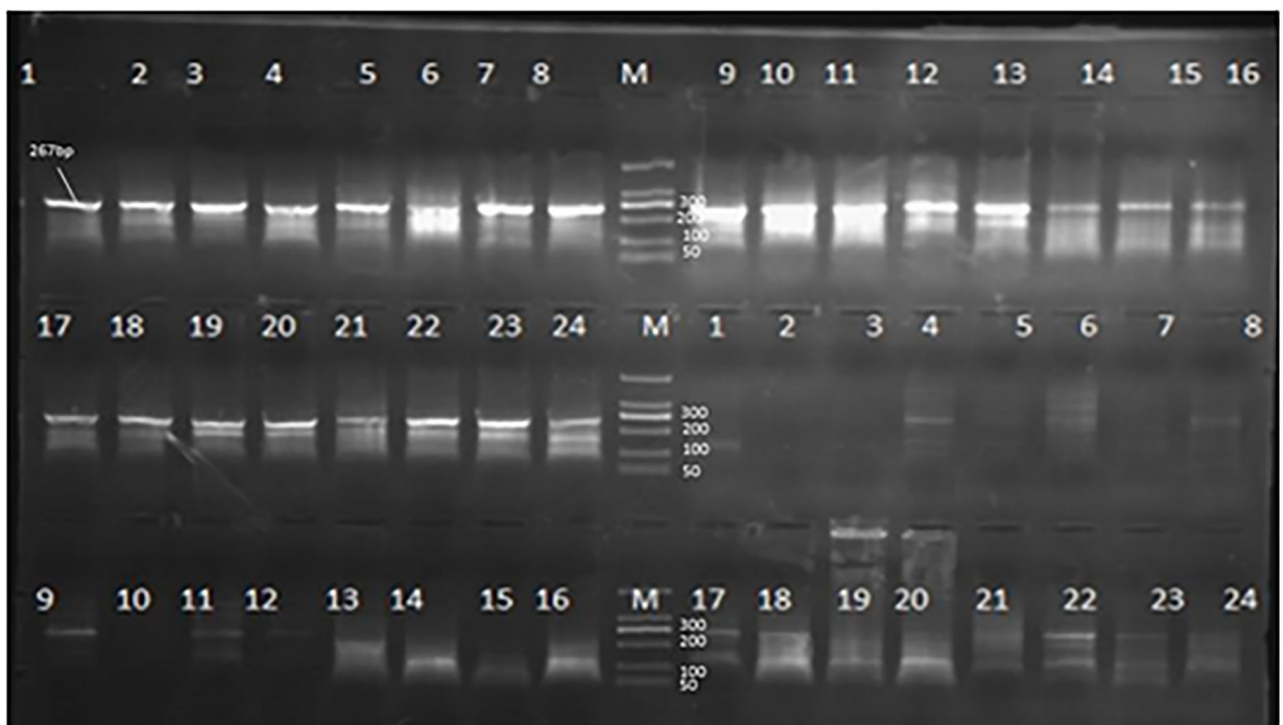


Fig 1. Depicted accession in study for C-T allele.

<https://doi.org/10.1371/journal.pone.0264478.g001>



Table 3. List of rice germplasm carrying C or T allele.

S. No	Land races	Allele	S. No	Land races	Allele
1	Rachna basmati	C	13	Line 134	C-T
2	Basmati 370	C	14	Line 48	C-T
3	IR9	C	15	Line 146	C-T
4	L-77	C	16	Basmati 385	C-T
5	KSK 434	C	17	NPT 22	C-T
6	Dilrosh	C-T	18	Basmati 200	C-T
7	Malkhar	C-T	19	KSK 133	C-T
8	Swat-I	C-T	20	P-5	C-T
9	Kashmir basmati	C-T	21	Muskan	C-T
10	L-20	C-T	22	Swat II	C-T
11	Line 46	C-T	23	Malakand	C-T
12	Line 31	C-T	24	P.S 2	C-T

<https://doi.org/10.1371/journal.pone.0264478.t003>

Line 77, KSK 434, Dilrosh, swat I, Kashmir basmati, Line 46, Line 48, Line 146, Basmati 385, NPT 22, Basmati 200, KSK 133, P-5, Muskan, Swat II, Malakand, P.S 2 carry heterozygous C-T Allele.

### *SPL14-12* C-A allele genotyping

PCR amplification was carried out for C-A allele in fragment size of 244bp and 302bp. All the accessions carry a C allele in fragment size of 244bp. C-A allele was located on third exon of *OsSPL14*, which is the target site of OsmiRNA156. OsmiRNA156 is a free miRNA that cleaves the mRNA transcript of the *OsSPL14* gene. Rachna basmati, Basmati 370, IR9, 77, KSK 434,

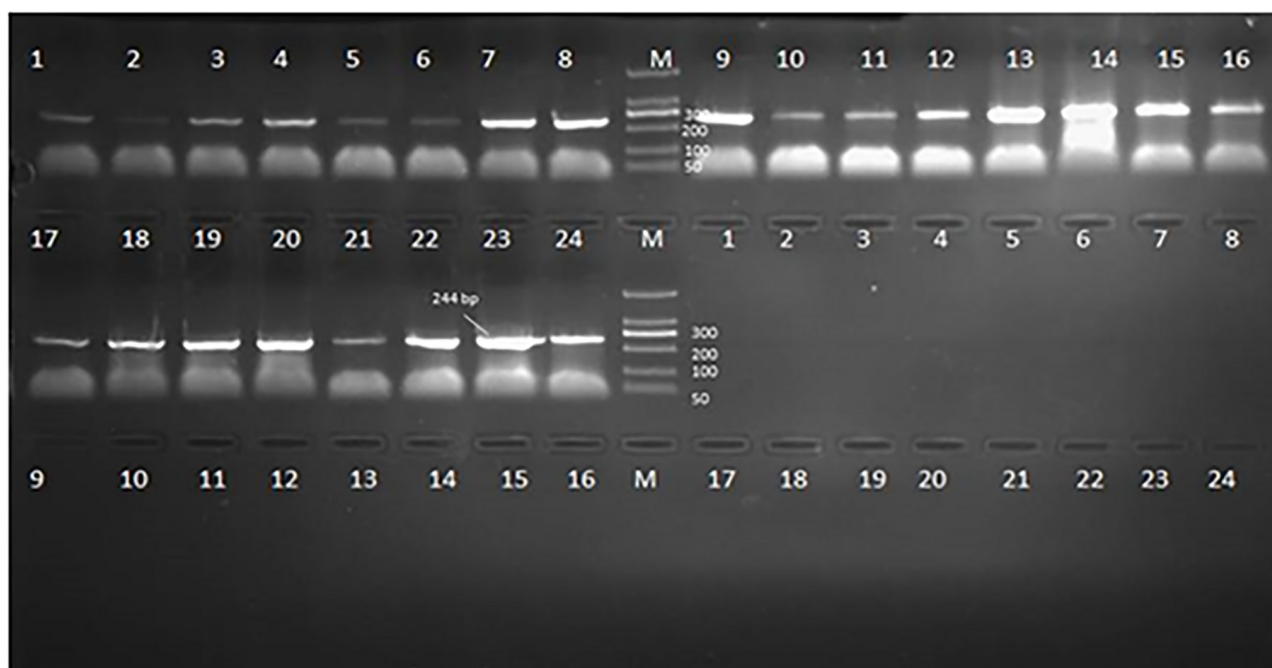


Fig 2. Depicted accession in study for C-A allele.

<https://doi.org/10.1371/journal.pone.0264478.g002>

Table 4. List of rice germplasm C- A allele.

S. No	Land races	Allele	S. No	Land races	Allele
1	Rachna Basmati	C	13	L-134	C
2	Basmati 370	C	14	L-48	C
3	IR9	C	15	L-146	C
4	LinE-77	C	16	Basmati 385	C
5	KSK-434	C	17	NPT 22	C
6	Dilrosh	C	18	Basmati 200	C
7	Malkhar	C	19	KSK 133	C
8	Swat-I	C	20	P-5	C
9	Kashmir Basmati	C	21	Muskan	C
10	L-20	C	22	Swat II	C
11	L-46	C	23	Fakher-e-Malakand	C
12	L-31	C	24	P.S 2	C

for **C allele** 1 = Rachna Basmati, 2 = Basmati 370, 3 = IR9, 4 = 77, 5 = KSK 434, 6 = Dilrosh, 7 = Malkhar, 8 = swat-I, 9 = Kashmir Basmati, 10 = Line 20, 11 = Line 46, 12 = Line31, 13 = Line 134, 14 = 48, 15 = Line 146, 16 = Basmati 385, 17 = NPT 22, 18 = Basmati 200, 19 = KSK 133, 20 = P.5, 21 = Muskan, 22 = Swat-II, 23 = Fakher-e-Malakand, 24 = P.S 2 M = Marker

<https://doi.org/10.1371/journal.pone.0264478.t004>

Dilrosh, Malkhar, Swat-I, Kashmir basmati, Line 20, Line 46, Line31, Line 134, Line 146, Basmati 385, NPT 22, Basmati 200, KSK 133, P.5, Muskan, Swat-II, P.S 2, DR 2 carried the wild type C allele (Fig 2, Table 4).

### Sequencing of *OsSPL14* for OSmiR156 target-site

The irregularity of visual correlation between tiller number per plant and the respective allele was sequenced at the 3<sup>rd</sup> exon of *OsSPL14* gene, which carries OSmiR156 directed transcriptional cleavage, if any additional mutation in miRNA target site. Interestingly, the sequence result revealed a new single nucleotide polymorphism (C-T substitution) in 09 different accessions, 2bp upstream of the previous reported SNP (C-A) [14, 18] in the target seed sequence for miRNA156 (Fig 3).

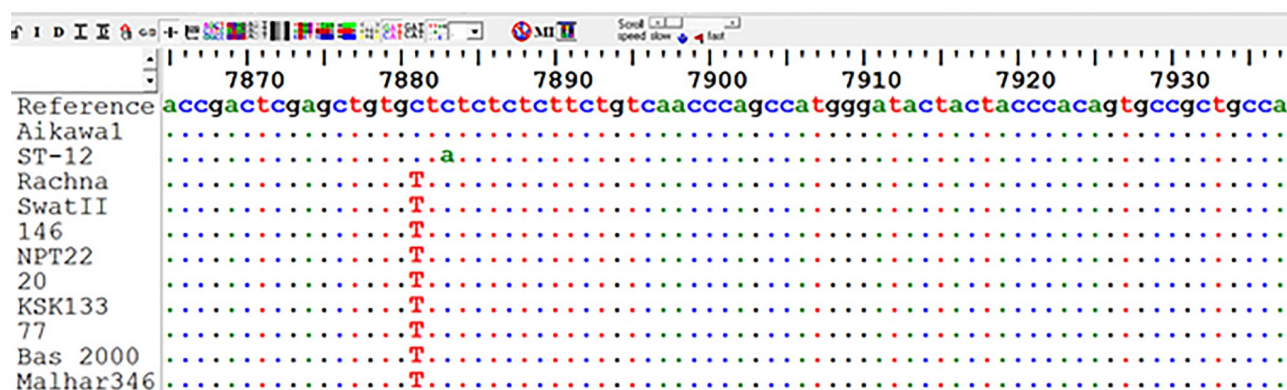


Fig 3. Aikawal and ST-12 are the two control sequences along reference genome (Nipponbare).

<https://doi.org/10.1371/journal.pone.0264478.g003>



Wild type	HGHSSAVAGYGANTYSGQGSSSSGPPVFAGPNLPPGGCLAGVGAATDSSCALSLSTQPW	300
<u>Valine</u>	HGHSSAVAGYGANTYSGQGSSSSGPPVFAGPNLPPGGCLAGVGAATDSSCALSLSTQPW	300
Wild type	DTTTHSAAASHNQAAAMSTTTSFDGNPVAPSAMAGSYMAPSPWTGSRGHEGGGRSVAHQ	360
<u>Valine</u>	DTTTHSAAASHNQAAAMSTTTSFDGNPVAPSAMAGSYMAPSPWTGSRGHEGGGRSVAHQ	360
Wild type	PHEVSLDEVHPGPSHHAHFSGELELALQNGPAPAPRIDPGSGSTFDQTSNTMDWSL	417
<u>Valine</u>	PHEVSLDEVHPGPSHHAHFSGELELALQNGPAPAPRIDPGSGSTFDQTSNTMDWSL	417

**Fig 4. Amino acid alignment to missense mutation indicates from alanine to valine.**

<https://doi.org/10.1371/journal.pone.0264478.g004>

## Analysis of protein for new mutation

The nucleotide sequence was translated using NCBI BLAST online tool which aligned translated sequence of new SNP (C-T) mutation carrying *OsPL14* revealed an amino acid substitution from alanine to valine (Fig 4). The aligned sequence showed 99% sequence homology with *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 14*. Both alanine and valine amino acids are internal, hydrophobic, and neutral.

To further understand the structural change in *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 14*, due novel mutation, the sequence results of both wild type and valine carrying gene were analyzed using I-TASSER that generated different models. The models chosen were based on RMSD, C-score, and TM-Score (Table 5). The best model was subjected to ERRAT2 software that predicted both sequences' quality factor (Fig 5).

The Models (Fig 6A & 6B) for wild type and alanine to valine substituted *OsSPL14* gene was selected among the five predicted models for each with large decoys (cluster density) based on pairwise structural similarities. C-score quantitatively measures each models' confidence based on the significance of threading template alignment and the structural assembly simulations' convergence parameter. C-score ranges from -5 to 2, which signifies high confidence of the model. A model of C-score more -1.5 indicated correct global topology. In the present study, the substituted amino acid alanine to valine of *OsSPL14* protein has a value of -0.79. Template Modelling Score (TM-score) compares two model structural alignment based on the amino acid residues and measures the structural homology. The TM-Score ranges from 0–1. The query substituted amino acid protein exhibited high TM-score (0.61+-14) compared to template Protein Data Bank in the present data. In Figs 5 and 6 at position 290 the substitution of amino acid from alanine to valine (Fig 5) shows the same error value as in the wild type (Fig 6A & 6B), and ERRAT2 generated the same quality factor score as for the wild type.

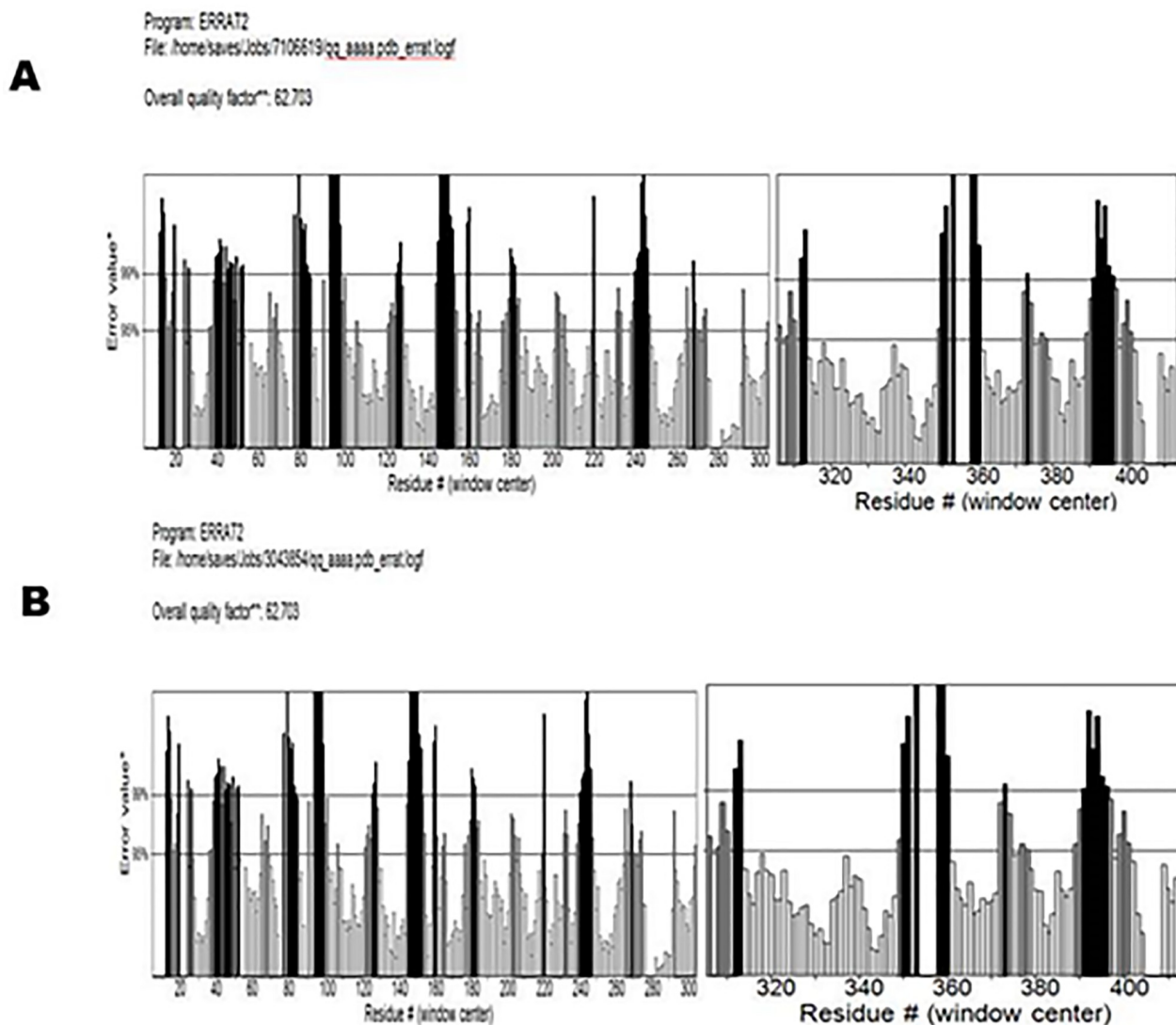
## Discussion

Rice tillers are produced on stunted basal internode that develop freely on the main stem having adventitious root. Tillering in rice is among the major constituent of rice grain yield [20].

**Table 5. Protein model score of the predicted wild and alanine to valine substitution.**

Models	Evaluation Scores			Overall quality factor (ERRAT2 used)
	C-score	Estimated RMSD	Estimated TM-Score	
Wild type <i>OsSPL14</i>	-1.08	9.4+-4.6A	0.58+-0.14	62.703
Alanine to Valine change in <i>OsSPL14</i>	-0.79	8.7+-4.5A	0.61+-0.14	62.703

<https://doi.org/10.1371/journal.pone.0264478.t005>

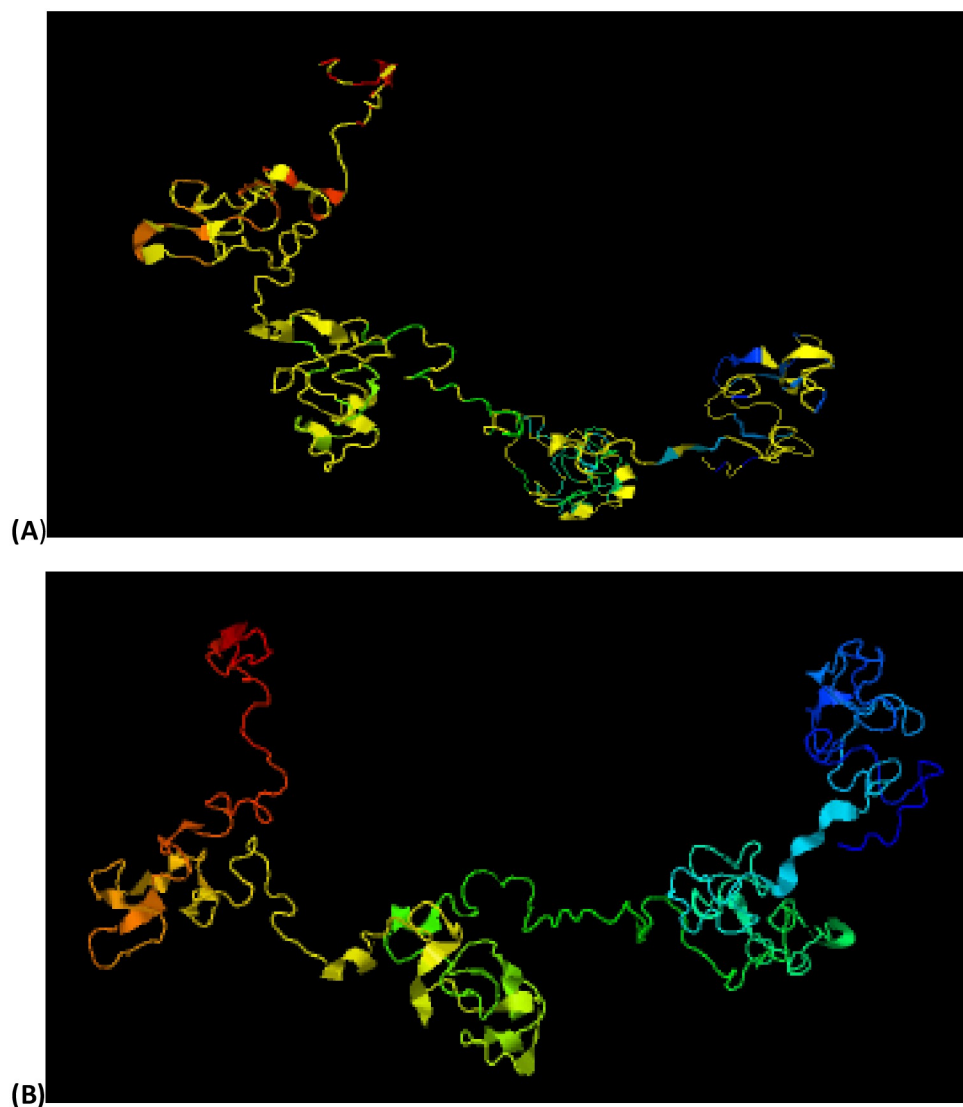


**Fig 5.** (A) Overall quality factor of wild type protein. (B) Overall quality factor of C-T substitution.

<https://doi.org/10.1371/journal.pone.0264478.g005>

*OsSPL14*, *IDEAL PLANT ARCHITECTURE* (*IPA*), and *WEALTHY FARMER GENE* (*WFP*) are major genes that tailors rice tillers and panicle branches. Regarding *OsSPL14* expression, *OsmiR156* plays an essential and dual role in tillers regulation, number of panicles branching and spikelet's development [14, 18]. This study analyzed 24 genotypes of rice accessions based on structural and molecular characterization to estimate the genetic diversity. Based on morphology, 6 traits were observed in 24 genotypes of rice. The sequence analysis showed presence of variation in the given data set among 24 genotypes i.e., Number of tillers plant, number of grains per panicle, secondary branched panicle, number of grains per plant, grain weight, as shown in (Table 1). In the present study number of tillers per plant, the number of grains per panicle, and the secondary branched panicle showed significant variation.

A highly significant correlation was observed in the number of grains per panicle with the number of grains per plant followed by primary branches per panicle and secondary branches per panicle, respectively. A significant correlation was recorded between the number of grains



**Fig 6. (A) Wild type (B) Alanine to Valine substitution.** Model generated by I-TASSER.

<https://doi.org/10.1371/journal.pone.0264478.g006>

per panicle and grain weight. Primary branches per panicle also showed a highly significant and strong correlation with secondary branches per panicle, followed by several grains per plant. The number of tillers per plant also showed a strong and highly significant correlation of 0.7620, 0.6313, and 0.5671 with primary branches per panicle, secondary branches per panicle, and grain weight, respectively. Similarly, the same work was conducted by [21], who observed a high variation among the seven traits. Correlation coefficients between all pairs of variables used in this experiment are shown in (Table 2).

The C allele ~2.6 kb mapped on the upstream at the start codon of SPL14-04 site is an epigenetic manner. The allele-specific amplification of this heritable epigenetic allele in the promoter region revealed nineteen samples as heterozygous (C and T). Previous research reported five heterozygous alleles out of 12 in BC<sub>2</sub>F<sub>2</sub> plants derived from a cross between PR38012 and ST-12 [17]. However, due to the heterozygosity of the allele in most of the accessions, no valid correlation was observed between the heritable epigenetic allele and the number

of tillers. Micro RNA plays an essential role in regulating many target genes. In plants, miRNA inhabits many genes, while in animals, it represses translation [22–24]. OsmiR156 expression lowered at the panicle development stage [14, 18]. The *SPL* genes are regulated by microRNA miR156, and *OsSPL16* of OsmiR156 was shown to repress *OsSPL16* transcription in young panicles [25].

In previous studies, no C-A mutation was found however, the sequence analysis result revealed a new C-T mutation located at ~2bp upstream of the *OsSPL14* gene. Hexapeptide ALSLLS is a conserved motif of *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 14* [26]. However, substitution with VLSLLS in the mutant type has no negative effect on the protein as predicted by I-TASSER and ERRAT2 model. The binding of miR156 to the seed sequence of *SPL-14* indicates the amount of *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 14* gene expressions. The average effect of the T allele in the seed sequence for miR156 on the number of grains per plant is the same as the wild-type allele, which suggests further insight that this new mutation (C-T) substitution did not inhibit the binding of the miR156 to the target site. Previous studies reported that microRNA imperfectly binds to seed sequence [25, 27, 28]. We also identified some lines that showed low grains per panicle, so we suggested that it might be some other genes regulating grain per panicle. Moreover, Line-20 and Line-77 carry low grains per panicle. Similarly, the grain yield per plant of Line-77, Line-20, and Line-48 also had low grain yield per plant.

## Conclusion

We identified for the first time a new C-T mutation in the third exon of the *OsSPL14* gene in rice which can play an important role in grain yield development. Mutation of *OsSPL14* in rice demonstrated the number of tillering and grain size based on phenotyping results, molecular characterization, and statistical analysis, respectively. Hence our present findings not only identified *OsSPL14* as a key regulator of grain yield but also suggest a strategy for the modification of grain yield in a wide range for the improvement in crops.

## Supporting information

**S1 Table. List of selected genotypes.**

(DOCX)

**S2 Table. Allele specific molecular marker used in the study.**

(DOCX)

**S1 Raw image.**

(DOCX)

## Acknowledgments

The authors extend their appreciation to the Researchers Supporting Project number (RSP-2021/241), King Saud University, Riyadh, Saudi Arabia.

## Permissions or licenses

Permissions or licenses were obtained. Experimental research and field studies on plants (either cultivated or wild), including collecting plant material, comply with relevant institutional, national, and international guidelines and legislation.

## Author Contributions

**Data curation:** Muhammad Abid Khan.

**Formal analysis:** Muhammad Salah ud Din, Salman Alamery, Kotb Attia, Itoh Kimiko.

**Funding acquisition:** Xiukang Wang, Sajid Fiaz.

**Investigation:** Sumbul Saeed.

**Methodology:** Niaz Ali.

**Software:** Shahid Ullah Khan.

**Writing – original draft:** Haroon Rasheed.

**Writing – review & editing:** Salman Alamery, Kalim Ullah Marwat, Kotb Attia, Itoh Kimiko, Shabir Hussain Wani.

## References

1. Qian Q., Guo L., Smith S. M. & Li J. Breeding high-yield superior quality hybrid super rice by rational design. *National Science Review* 2016, 3, 283–294.
2. Liang W. hong, Shang F., Lin Q. ting, Lou C. & Zhang J. Tillering and panicle branching genes in rice. *Gene* 2014, 537 1–5. <https://doi.org/10.1016/j.gene.2013.11.058> PMID: 24345551
3. Fujita D. et al. Fine mapping of a gene for low-tiller number, *Ltn*, in japonica rice (*Oryza sativa* L.) variety Aikawa 1. *Theor. Appl. Genet.* 2010, 120, 1233–1240. <https://doi.org/10.1007/s00122-009-1251-7> PMID: 20062964
4. Iqbal M. M. et al. Climate Change and Rice Production in Pakistan: Calibration, Validation and Application of CERES-Rice Model. in Islamabad: Global Change Impact Studies Centre (2009).
5. Abedullah S. K. & Mushtaq K. Analysis of technical efficiency of rice production in Punjab (Pakistan). *Pak. Econ. Soc. Rev* 2007, 45, 231–244.
6. Von Braun, J. & Bos, M. S. *The changing economics and politics of rice: implications for food security, globalization, and environmental sustainability. IRRI reports* (2005).
7. Jeon J. S., Jung K. H., Kim H. B., Suh J. P. & Khush G. S. Genetic and Molecular Insights into the Enhancement of Rice Yield Potential. *Journal of Plant Biology* 2011, 54 1–9.
8. Aleksieva, I. A. & Szeker, D. Rice, Monsoon and Revolution: The past, present and future of rice production in India, a project on the effects of the Green Revolution and Climate Change on the Indian rice cultivation. (2019).
9. Muthayya S. et al. Rice fortification: An emerging opportunity to contribute to the elimination of vitamin and mineral deficiency worldwide. *Food Nutr. Bull* 2012, 33, 296–307. <https://doi.org/10.1177/156482651203300410> PMID: 23424896
10. Xing Y. & Zhang Q. Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol* 2010, 61, 421–442. <https://doi.org/10.1146/annurev-arplant-042809-112209> PMID: 20192739
11. Gomez-Roldan V. et al. Strigolactone inhibition of shoot branching. *Nature* 2008, 455, 189–194. <https://doi.org/10.1038/nature07271> PMID: 18690209
12. Umehara M. et al. Inhibition of shoot branching by new terpenoid plant hormones. *Nature* 2008, 455, 195–200. <https://doi.org/10.1038/nature07272> PMID: 18690207
13. Miura K. et al. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet* 2010, 42, 545–549. <https://doi.org/10.1038/ng.592> PMID: 20495564
14. Kakutani T. Epi-alleles in plants: Inheritance of epigenetic information over generations. *Plant Cell Physiol.* 2002, 43, 1106–1111. <https://doi.org/10.1093/pcp/pcf131> PMID: 12407189
15. Kim S. R. et al. Introgression of a functional epigenetic *OsSPL14* WFP allele into elite indica rice genomes greatly improved panicle traits and grain yield. *Sci. Rep* 2018, 8, 3833. <https://doi.org/10.1038/s41598-018-21355-4> PMID: 29497052
16. Jiao Y. et al. Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. *Nat. Genet.* 2010, 42, 541–544. <https://doi.org/10.1038/ng.591> PMID: 20495565
17. Endo-Higashi N. & Izawa T. Flowering time genes heading date 1 and early heading date 1 together control panicle development in rice. *Plant Cell Physiol.* 2011, 52, 1083–1094. <https://doi.org/10.1093/pcp/pcf059> PMID: 21565907



18. Chang, T. & Bardenas, E. The morphology and varietal characteristics of the rice plant. The International Rice Research Institute. Technical Bulletin (The International Rice Research Institute, 1965).
19. Steel R. G., Torrie J. H. & Dickey D. A. Principles and Procedures of Statistics: A Biometrical Approach. ( McGraw Hill Book International Co., 1997).
20. Li X. et al. Control of tillering in rice. *Nature* 2003, 422, 618–621. <https://doi.org/10.1038/nature01518> PMID: 12687001
21. Yu H. et al. Fine mapping of *LOW TILLER 1*, a gene controlling tillering and panicle branching in rice. *Plant Growth Regul.* 2017, 83, 93–104.
22. Perween S. Molecular characterisation of genetic diversity in aromatic rice genotypes of chattisgarh region. *Journal of Crop and Weed* 2017, 14, 94–99.
23. Wang S. et al. The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 2015, 47, 949–954. <https://doi.org/10.1038/ng.3352> PMID: 26147620
24. Stefani G. & Slack F. J. Small non-coding RNAs in animal development. *Nature Reviews Molecular Cell Biology* 2008, 9, 219–230. <https://doi.org/10.1038/nrm2347> PMID: 18270516
25. Mallory A. C. & Vaucheret H. Functions of micromnas and related small a mas in plants. *Nat. Genet.* 2006, 38, S31–S36. <https://doi.org/10.1038/ng1791> PMID: 16736022
26. Jones-Rhoades M. W. Conservation and divergence in plant microRNAs. *Plant Molecular Biology* 2012, 80, 3–16. <https://doi.org/10.1007/s11103-011-9829-2> PMID: 21996939
27. Xie K., Wu C. & Xiong L. Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* 2006, 142, 280–293. <https://doi.org/10.1104/pp.106.084475> PMID: 16861571
28. Bari A., Orazova S. & Ivashchenko A. Mir156- and mir171-binding sites in the protein-coding sequences of several plant genes. *Biomed Res. Int.* 2013, 307145. <https://doi.org/10.1155/2013/307145> PMID: 23936788